

1990

Reasons, Authority, and the Meaning of 'Obey': Further Thoughts on Raz and Obedience to Law

Donald H. Regan

University of Michigan Law School, donregan@umich.edu

Available at: <https://repository.law.umich.edu/articles/341>

Follow this and additional works at: <https://repository.law.umich.edu/articles>



Part of the [Jurisprudence Commons](#), and the [Law and Philosophy Commons](#)

Recommended Citation

Regan, Donald H. "Reasons, Authority, and the Meaning of 'Obey': Further Thoughts on Raz and Obedience to Law." *Canadian J. L. & Juris.* 3 (1990): 3-28.

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

Reasons, Authority, and the Meaning of "Obey": Further Thoughts on Raz and Obedience to Law

Donald H. Regan

I recently published a long article¹ discussing a variety of topics from Joseph Raz's *The Morality of Freedom*.² The article was part of a symposium on Raz's work in the *Southern California Law Review*. Raz responded³ to the articles in that symposium, including my own.

From a perspective which surveys the whole range of views on political philosophy, Raz's view and mine look very similar. Even so, we find many things to disagree about, which neither of us would regard as merely matters of detail. For the most part, we at least share a common understanding of our disagreements. But there is one set of issues we disagree about and where we seem to lack even a common understanding of the disagreement. These are issues about how authoritative directives function as reasons for action, and about whether, when we follow authority in the proper way, we can be said to "obey" it.

Raz says he finds "mystifying" my distinguishing between (in his words) "true moral reasons" and "the sort of considerations that . . . validate claims of . . . obedience".⁴ I take it what Raz is referring to is my distinguishing between what I call "intrinsic" and "indicative" reasons for action.⁵ If I have not made the content and the importance of this distinction sufficiently clear, it is worth trying again. In this essay, I shall try to separate the core distinction from auxiliary issues that may have caused confusion in my earlier treatment. I shall also be more explicit about why the distinction is important.

In addition, I shall discuss more explicitly than before what I think is involved in obedience. I do not claim, as Raz suggests, that indicative reasons but not intrinsic reasons justify us in obeying authority. (It would never occur to me to

This essay was largely written at the National Humanities Center, with the support of the Andrew W. Mellon Foundation. My thanks to both. I am also grudgingly grateful to Dick Brunaugh, whose helpful comments increased the length of the essay by a third.

1. Donald H. Regan, "Authority and Value: Reflections on Raz's *Morality of Freedom*" (1989), 62 *Southern California Law Review*, 995-1095.
2. J. Raz, *The Morality of Freedom*, (Oxford: Clarendon Press, 1986).
3. Raz, "Facing Up: A Reply" (1989), 62 *Southern California Law Review*, 995-1235.
4. *Ibid.* at 1197, n. 91.
5. I introduced this terminology in "Law's Halo" (1986), 4 *Social Philosophy & Policy* at 15-30. I referred to this terminology in "Authority and Value" (1008 n.40, 1022 n.65), but I did not use it in "Authority and Value", which I now regret. In "Authority and Value", trying to keep as close as possible to Raz's terminology, I referred instead to "reasons for action in the strict sense" (intrinsic reasons) and "reasons for action in the loose sense" (indicative reasons). As I shall explain later in this essay, I think Raz's terminology is problematic, and I would have been clearer if I had distanced myself from it more completely. Part of my motive was that I was concentrating in "Authority and Value" on the concept of an "indicator-rule", which we might think of as a special use we can make of certain unusually reliable indicative reasons. So, an indicator-rule is not just an indicative reason. But the concept of an indicator-rule cannot be properly understood unless we already have the concept of an indicative reason. Since some expansion on the discussion of indicative reasons in "Law's Halo" was undoubtedly called for, the present essay really should have been written before "Authority and Value". We often realize these things too late.

make a claim of this form, since indicative reasons get whatever force they have by indicating the putative existence of some intrinsic reason or other.) My claim about obedience is that *no* sort of reason actually justifies us in *obeying* authority, in the proper philosophical sense of “obey”, even though authoritative directives often matter to what we ought to do. The right relationship to authority (or at least, to authority justified as Raz justifies it) is never one of obedience, properly understood. Incidentally, we shall see that a proper understanding of obedience depends on the distinction between intrinsic and indicative reasons, so everything comes together in the end.

I suspect one reason Raz and I do not yet understand one another is that we have been distracted from the most fundamental issue about how authoritative directives function as reasons by controversies about the exclusionary force of authoritative directives and about the promotion of coordination. These controversies do involve the most fundamental issue, but unfortunately they involve much else besides. In this essay I shall concentrate on how authoritative directives function as *first-order* reasons in cases where coordination is not the problem. I shall not discuss, except in passing, authoritative directives as exclusionary reasons (that is, as second-order reasons for not acting on or not considering other reasons). Nor shall I discuss how authoritative directives promote coordination.

So much of the criticism of Raz’s accounts of reasons and of authority has focussed either on the nature of exclusionary reasons or on coordination, that it may be hard to credit that we have a disagreement on the seemingly elementary topic I propose to discuss. Even my own article on Raz may have encouraged the idea that all my concerns were with exclusionary force or with coordination. But in fact, there is a problem about how authoritative directives function as first-order reasons in non-coordination cases. And I think this is the case that presents the fundamental issue about authoritative directives in its simplest form.

The fundamental issue is whether authoritative directives are intrinsic reasons or indicative reasons. As it happens, we can distinguish between intrinsic and indicative reasons both within the class of first-order reasons and within the class of second-order (in particular, exclusionary) reasons. And I think much of the debate about the exclusionary force of authoritative directives is best understood as being about whether this exclusionary force is intrinsic or indicative. But that debate is confused by the fact that the corresponding issue about the *first-order* force of authoritative directives is also implicated but goes unnoticed.

I think the best hope for settling the debate about exclusionary force, and similarly for settling the debate about how authority helps with coordination, is to begin by ignoring exclusionary force and coordination. If we can clarify how authoritative directives function as first-order reasons in cases where coordination is not a problem, and in the process get a firm grip on the intrinsic/indicative distinction and its significance, the issues about exclusionary force and coordination ought to unravel themselves fairly easily.⁶

So, what is the distinction between intrinsic and indicative reasons? Roughly

6. Having said I shall ignore coordination, I must insert a footnote about it. In his response to me and to Leslie Green, “Law, Legitimacy, and Consent” (1989), 62 *Southern California Law Review*, 795 at 804-8, Raz adds something important to the discussion of coordination in his book. (“Facing Up”, *supra* note 3 at 1187-94) He points out that one is not “in” a coordination problem as game theorists normally understand it

speaking, an intrinsic reason is a fact about a proposed action that matters *in itself* to practical reasoning, while an indicative reason is a fact that matters because it is *evidence* for the existence (or non-existence) of some fact that constitutes an intrinsic reason. Notice that some fact could be both an intrinsic reason and an indicative reason, if it both mattered in itself and also was evidence for the existence of some other fact which mattered in itself. But for the most part I shall use the phrase "indicative reason" to mean a reason which is *merely* indicative (that is, which does not matter in itself but only as evidence), so that the categories "intrinsic reason" and "indicative reason" are mutually exclusive.

I shall expand on this rough statement of the distinction between intrinsic and indicative reasons, and give some examples, in a moment. But let me digress briefly to explain how the distinction underlies the disagreement between Raz and me about the significance of authoritative directives, and also to say something about Raz's terminology for what may or may not be the same distinction. What I have to say about intrinsic and indicative reasons will be more meaningful for the reader if she has in mind some of the specifics of my disagreement with Raz.

The central disagreement between Raz and me is this: In my view, authoritative directives are (merely) indicative reasons; for Raz, so far as I can see, authoritative directives are intrinsic reasons. (They are indicative reasons also, but they are not merely indicative.) Raz does not say that authoritative directives are intrinsic reasons in so many words, since he does not use this terminology. Indeed, certain aspects of his view ought to make him want to deny that authoritative directives are intrinsic reasons. But I think he is committed to the view that authoritative directives are intrinsic reasons by his claim that when we follow authority we *obey* it. Raz explicitly connects up the notion of obedience with authoritative reasons' being "reasons for action";⁷ and I shall explain later, when I look more closely at the meaning of obedience, why the claim that we obey authority requires that authoritative directives should be *intrinsic* reasons. (Notice I have said that certain aspects of Raz's view suggest that authoritative directives are merely indicative reasons. So in a sense my criticism of Raz is not that he single-mindedly believes authoritative directives are intrinsic reasons, but rather that he holds two incompatible views without acknowledging the tension. More on this later.)

unless everyone involved knows the basic situation, and is motivated to coordinate, and knows that others know the situation and that they are motivated to coordinate, and so on. And so there is a sense in which just *recognizing* a coordination problem is itself a coordination problem. Each would-be coordinator must correctly identify the others, and the identification must be known to be mutual. (For practical purposes in many-person cases the identification may not need to be perfect, but that does not affect the principle.) Raz now suggests that a central aspect of government's contribution is that it helps us not merely to solve first-order coordination problems, but also to solve the second-order problem of recognizing the first-order ones. About this, I think he is absolutely right. (I cannot resist saying that although I had not previously noticed this point about the function of government, I had noticed that recognizing a coordination problem is itself a coordination problem at a higher level. The repellent complexity of the "full-dress" decision procedure for cooperative utilitarianism that I presented in *Utilitarianism and Co-operation* (Oxford: Clarendon Press, 1980) is entirely attributable to the need to solve the higher-order problem.) However, the realization that government helps us to identify coordination problems as well as to solve them, while it complicates a full description of what goes on, does not convert authoritative utterances into intrinsic reasons. Raz virtually admits this when he says that if some coordinative practice government tries to institute does not catch on, then individuals are justified in ignoring it. ("Facing Up", *supra* note 3 at 1194)

7. Raz, *supra* note 2 at 29.

Incidentally, it is because I think authoritative directives are not intrinsic reasons that I deny that we ever properly “obey” authority in the fullest sense. Raz says that if I want to sever the connection between authority and obedience, then I cannot be analyzing the ordinary concept of authority.⁸ That is probably correct. In any event, I do not mean to be analyzing the ordinary concept of authority. My object is to describe the proper relationship between what we ordinarily speak of as “authorities” and “subjects”. I would not use the word “authority” at all except that Raz uses it and I am responding to Raz. Once I admit that I am not trying to analyze the ordinary concept of authority, Raz might say that he and I are engaged in different projects, since he does want to analyze the ordinary concept. But I shall claim later that Raz’s analysis of the ordinary concept contains significant internal tensions, if not downright inconsistencies. This means that either Raz has not successfully analyzed the ordinary concept, or else the ordinary concept is incoherent in ways Raz overlooks.

I have mentioned more than once that Raz has a different terminology from mine, and that the relationship between his terminology and mine is problematic. One distinction Raz makes is between “reasons for belief” and “reasons for action”; and some passages in Raz seem to make best sense if we identify his “reasons for belief” with my “indicative reasons” and his “reasons for action” with my “intrinsic reasons”.⁹ But if this is what Raz intends by his terms, then his terms are misleading, because (as we shall see) *both* indicative reasons and intrinsic reasons are “reasons for action” in the most natural sense of that phrase; they are both what Raz elsewhere calls “guiding reasons”.¹⁰ They are both facts about the world that (ought to) guide action, as opposed to facts about an agent’s beliefs which merely explain action. I think Raz is sometimes led by his terminology of “reasons for belief” versus “reasons for action” to treat indicative reasons as mere explanatory reasons, with no proper significance for guiding action. This is a mistake. It is probably not a mistake Raz would commit consciously and explicitly. But something like this seems to lie behind Raz’s complaint that Stephen Perry, whose criticisms of Raz overlap with mine, “is inclined to think that uncertainty can be accounted for only by regarding people’s beliefs . . . as [guiding] reasons.”¹¹ This does not seem to me the best reading of Perry, whose real point is that in situations of uncertainty, what *facts about the world* are properly treated as guiding reasons (in particular, as indicative reasons) is partly determined by the general state of the agent’s knowledge and belief relevant to the decision at hand.¹²

8. “Facing Up”, *supra* note 3 at 1186.

9. For example, Raz, *supra* note 2 at 29. In one sense, of course, “reasons for belief” is definitely broader than “indicative reasons”, since we may have reasons for belief about facts which have no bearing on action. But if we agree that we are at all times talking in a practical context (as Raz normally is), so that the beliefs we are interested in are beliefs which do have a bearing on action, then the identification suggested in the text is plausible.

10. “Facing Up”, *supra* note 3 at 1154, n. 2.

11. *Ibid.*

12. Stephen Perry, “Second-Order Reasons, Uncertainty and Legal Theory” (1989), 62 *Southern California Law Review* 913-94, at 919-27. Perry may invite misunderstanding by an occasional comment and by his terminology of “subjective” and “objective” exclusionary reasons. But Perry makes some very astute observations. He is quite right that practical reasoning must sometimes deal in probabilities, and that it is far from obvious *precisely* how we should describe reasoning from probabilities within the context of the general model (char-

Intrinsic and Indicative Reasons

Let me now expound more fully the distinction between intrinsic and indicative reasons. I shall focus on first-order reasons, and not everything I say can be applied without alteration to second-order reasons, but the basic idea is the same at all levels. An *intrinsic* reason, as I have said before, is a fact about a proposed action that matters *in itself*; it is a fact that counts in itself in favor of or against some course of action. Another way to put this is that intrinsic reasons are facts that count *in the final analysis*: they would have a place in the deliberation of an agent who was fully informed, perfectly competent at whatever weighing and balancing of reasons is required, properly motivated, free of painful anxiety surrounding decision-making, and so on. (This latest formulation may seem to commit me to agent-neutrality, that is, to the belief that in the final analysis, exactly the same facts are intrinsic reasons, and have the same weight as intrinsic reasons, for all moral agents. As it happens, I accept agent-neutrality, but the distinction between intrinsic and indicative reasons does not depend on assuming neutrality. Even if different things matter in themselves for the deliberation of different agents, we can still distinguish, for each agent, between those things that matter in themselves – that would matter if that agent were fully informed and so on – and those things that matter only derivatively.)

An indicative reason, in contrast, is something that matters as *evidence* for the existence of an intrinsic reason. I have pointed out earlier that one and the same fact might be both an intrinsic reason and an indicative reason (if it both matters in itself and is evidence for some other fact that matters in itself). But if we use the phrase "indicative reason" to mean something which is *merely* an indicative reason, then we can say that indicative reasons do not matter in themselves. When they matter, they matter only as evidence for the existence of intrinsic reasons. It follows that they cease to matter, not only when an agent is fully informed, but even when an imperfectly informed agent has better evidence available on the relevant issue.

Let us have an example. Suppose a canvasser turns up at my door and requests a contribution to the Fund for the Relief of Victims of Hurricane Hugo (which I know to be a legitimate organization currently soliciting contributions in my neighborhood). The canvasser presents credentials which indicate that he represents that organization. I may reason as follows: "This person has credentials that are good evidence that he represents the Fund for the Relief of Victims of Hurricane Hugo. Therefore, if I give him money, it will relieve some suffering. Therefore, I shall give him money." Now, in this bit of reasoning, there is reference both to an intrinsic reason and to an indicative reason. The intrinsic reason for giving money to the canvasser is the fact that "if I give him money, it will relieve some suffering." The relief of suffering, or more expansively, the fact that some

acterized by respect for the distinction between beliefs-as-explanatory-reasons and facts-as-guiding-reasons) which Perry is as committed to as Raz and I are. Perry also points out a striking difference between the way authoritative directives function as exclusionary reasons and the way promises function as exclusionary reasons (both according to Raz). Authoritative directives have their exclusionary force because they *represent* certain other reasons; promises do not represent other reasons, but just make certain sorts of reason straightforwardly irrelevant. Perry does well to ask whether reasons whose exclusionary force is so differently grounded are likely to operate exactly the same way in deliberation.

possible act will relieve suffering, is what matters in itself. But what leads me to believe in the existence of this fact that constitutes an intrinsic reason is an *indicative* reason, namely, the fact that the canvasser has plausible-looking credentials. The possession of credentials does not matter in itself. It matters, but it matters only because the canvasser's possession of these credentials indicates that if I give the canvasser money, some suffering will be relieved.

To put the difference between intrinsic and indicative reasons in greater relief, let us compare different ways in which a more complicated deliberation might lead me not to give the canvasser the money after all. One possibility is that I am very poor, and that I in fact need what money I have even more than the victims of Hugo need it. If this is the case, I can properly decide not to make the requested contribution. In deciding not to make the contribution, I do not in the least depreciate the intrinsic reason which I have for giving the canvasser money, namely, that it will relieve suffering. But I recognize a competing and stronger intrinsic reason for not giving him the money, namely, that it will relieve even more suffering, of my own and my family's, if I keep the money and put it to my own uses. The intrinsic reason I have for giving the money to the victims of Hugo is *outweighed* by a stronger reason pointing in the other direction.

The second possibility is this. Suppose that, just as I am reaching for my wallet, my neighbor rushes up to tell me she has just learned the canvasser is an impostor who is lining his own pocket. His credentials are a clever forgery. Once again, assuming I believe my neighbor in preference to the canvasser, I will not give the canvasser any money. But that will not be because the intrinsic reason for giving him money is outweighed. Rather, it will be because I see there is in fact no intrinsic reason to give him money at all. It is not the case that giving him money will relieve suffering. The original indicative reason which was evidence for the proposition that giving the canvasser money would relieve suffering was misleading. (This makes it clear, incidentally, that I have not been using the verb "indicate", or such related phrases as "be evidence for", as success-verbs. Something may be an indicator, even a generally very good indicator, and still on occasion mislead.) Once it is revealed that the canvasser is an impostor, the force of the original indicative reason as a guide to action vanishes. The intrinsic reason that the indicative reason (misleadingly) indicated is not outweighed; rather, it drops out of the calculus completely.

In sum, intrinsic reasons, assuming they exist, matter in themselves. They may be outweighed, but, because they matter in themselves, they do not thereby lose their relevance to the final calculation. They are just outweighed. Indicative reasons, in contrast, do not matter in themselves. They matter only as indicators. Because they do not matter in themselves, they may be totally annulled, so far as their relevance to the final calculation is concerned. This happens if we decide, on the basis of other more reliable indicators, that the intrinsic reasons they point to do not really exist after all.

(Notice I have not meant to claim, in this discussion of why I might not give to the Hugo canvasser, that intrinsic reasons can *never* be subject to any more drastic fate than being outweighed; I have not meant to claim they can never be totally deprived of relevance. To say that would be to deny the possibility of exclusionary reasons as understood by Raz, which do make some intrinsic reasons

irrelevant. As I have said before, I am simply not focussing in this essay on exclusionary force. But the example of the Hugo canvasser involves no exclusionary reasons. So long as we limit ourselves to cases that involve no exclusionary reasons, we can say that intrinsic reasons, if they exist, never lose their relevance but are only susceptible of being outweighed, whereas indicative reasons may lose their relevance entirely. (Notice conversely that there is a certain perspective from which we can say that even an indicative reason is merely "outweighed". We might say the canvasser's credentials were outweighed *as evidence* by my neighbor's information. So, they were outweighed at the stage of my deliberation where I was considering competing evidence for and against the existence of the intrinsic reason in question, that giving the canvasser money would relieve suffering. But once I had decided against the canvasser's credentials at that stage, their relevance to the ultimate decision was at an end. The intrinsic reason whose existence they putatively indicated was not outweighed; rather, it evaporated.)

At the risk of introducing an unnecessary complication, let me mention a point that may be nagging at some readers. In the most general analysis, we would see that there are two different ways in which indicative reasons can operate, corresponding to two different kinds of proposition they may indicate the truth of. In the course of making any moral decision, we rely, whether consciously or not, on two kinds of proposition. On the one hand, we need to know propositions about what facts would count as intrinsic reasons for and against the various actions we are considering if those facts actually existed; and on the other hand, we need to know which of the potentially relevant facts do exist. Thus, in deciding that I ought to do a particular act which would relieve suffering, I rely both on the proposition (a) that the fact that some act would relieve suffering is a reason to do that act, and on the proposition (b) that the particular act in question would actually relieve suffering. (If the decision is a more complicated one, I may have to take into account other such propositions as well, but always in pairs, as it were: one proposition about what facts count as reasons if they exist, and another proposition about what the facts are.)

Now, in the example of the Hugo canvasser, the indicative reason I have identified indicates a proposition about the existence of a relevant fact. The canvasser's credentials indicate that giving money to the canvasser will relieve suffering. But there can also be indicative reasons which lend support to the other sort of proposition, propositions about what facts, if they exist, count as reasons. Suppose I am wondering whether it is wrong to commit suicide just to escape pain (that is, whether the fact that an act would amount to suicide-for-the-purpose-of-escaping-pain is a reason against performing the act). It occurs to me that Kant thought suicide to escape pain was wrong. Now, Kant was by all accounts a highly moral, if mildly eccentric, man in his personal life; and he was also one of the greatest of moral philosophers. I could quite reasonably say to myself that Kant's thinking such suicide was wrong gives me also some reason to think it is wrong. (This, incidentally, is regarding Kant as a "moral authority" in the most ordinary sense of that phrase.) But of course, Kant's believing that suicide to escape pain is wrong is only an *indicator* of the wrongness of such suicide; it is not *conclusive*. Even Homer nodded; even Kant may have erred on certain issues (as most modern moral philosophers think he did on this issue about suicide).

Because there are differences in the way we know such propositions as “this act would relieve suffering” and “that some act would relieve suffering is a reason for doing it”, indicative reasons which indicate these different sorts of proposition operate somewhat differently. However, I do not propose to go into that. Most of what I have to say about indicative reasons will apply equally to indicative reasons of either sort. Insofar as it matters, the indicative reasons I shall have chiefly in mind are reasons which indicate the existence of facts (that is, reasons which indicate propositions like “this act would relieve suffering”). These are the sort of indicative reasons about which I intend that everything I say in this essay shall be true.

Now, back to the main business. I said some time ago that indicative reasons, quite as much as intrinsic reasons, were “guiding reasons” in Raz’s sense. They are facts that guide our actions, and not merely beliefs that explain them. Remember the summary of my reasoning when the canvasser came to my door and I decided to give him money: “This person has credentials that are good evidence that he represents the Fund for the Relief of Victims of Hurricane Hugo. Therefore, if I give him money, it will relieve some suffering. Therefore, I shall give him money.” The indicative reason here – the fact that the canvasser has the credentials – is a fact, not a mere belief on my part. And this fact guides my action.

It is true that the indicative reason guides my action indirectly, and guides it by guiding my belief about an intrinsic reason. The credentials lead me to believe that giving the canvasser money will relieve suffering; and it is that latter belief (or that fact, as I now believe it to be) that I rely on directly when I actually give him the money. But even indirect guidance is guidance. Also, the fact that a belief has entered the picture does not mean we are now dealing in explanation. *Whenever* we act, we rely on our beliefs about the existence of intrinsic reasons in the way I rely on my belief that my act will relieve suffering in this example. (In an obvious sense, our only access to the facts is through our beliefs.) And in almost all cases, our beliefs about intrinsic reasons are guided by indicative reasons, just as in this example. We rarely if ever have direct, unmediated knowledge of all the relevant intrinsic reasons. Indicative reasons are not only guides to action, they are often the only guides we have immediate knowledge of.

Perhaps it will help to see that indicative reasons are guides to action, even though what they *directly* ground is beliefs, if I compare the way beliefs figure in the example just given with the way beliefs figure in an explanation of behavior. Suppose I have given the canvasser money, and then my neighbor appears. She saw through the canvasser’s imposture immediately, and assumes that I would have also. When I tell her I gave the canvasser money, she asks me why. I respond, “Because I thought he was genuine. I thought the money would go to victims of Hugo.” Here I am explaining my action. (There is also a sense in which I am trying to “justify” myself, but “justifying” an action in this sense really means *explaining* it in such a way that the agent does not seem a knave or a fool.)

The crucial distinction between the context of explanation and the context of action-guiding is this: When we are dealing in explanation, the beliefs are primary, the facts secondary; whereas, when we are engaged in action, the facts are primary, the beliefs secondary. When we are engaged in action, we are trying to make

our beliefs about the reasons fit the facts. We must be relying on our beliefs, because our only access to the facts is through our beliefs. But we do not rely on them essentially as "our beliefs". That is why we can give a summary of our reasoning, even from indicative reasons, which, like the summary I gave in the case of the canvasser, makes no explicit reference to our beliefs at all. (It is true, as I have mentioned previously, that what counts as an indicative reason depends on the state of the agent's relevant general knowledge and belief. That is why there was a sort of implicit reference to my beliefs when, in the summary of my reasoning about the Hugo canvasser, I referred to his credentials as "good evidence". But even here, the generalizations about the world that I rely on, usually unconsciously, in deciding what facts count as indicative reasons, I rely on not as "my beliefs", but rather as true generalizations about what facts tend to co-occur.)

In contrast, when we are *explaining* some action, the agent's beliefs are primary. The agent was presumably trying to make her beliefs fit the facts; and in the absence of contrary evidence, we may therefore attribute to her beliefs that truly reflected the facts as we know them. But even so, if we are dealing in explanation, it is the agent's beliefs that matter essentially. If her beliefs did *not* fit the facts, then it is the facts that disappear from the explanation, not the beliefs.

I hope it is now clear that indicative reasons are guiding reasons. *They figure in a process, the process of making decisions, in which the facts are primary.* They guide our actions by guiding our beliefs about the presence or absence of intrinsic reasons, but that does not reduce them to mere beliefs-as-elements-of-an-explanation.

Why the Intrinsic/Indicative Distinction Matters (In Brief)

I have now explained the distinction between intrinsic and indicative reasons, and I have attempted to lay to rest any doubts about whether indicative reasons are guiding reasons. It seems to me it should be obvious why the distinction between intrinsic and indicative reasons is important, but in case it is not, let me try to make the case quickly, partly by highlighting points we have already encountered, and partly by a further observation.

As to the points we have already encountered: First, even though both indicative and intrinsic reasons are indispensable to most exercises in practical reasoning, there is a sense in which intrinsic reasons are logically primary and indicative reasons logically secondary. Because indicative reasons have their force as indicators of the existence of intrinsic reasons, we cannot tell what counts as an indicative reason for a particular agent without knowing what facts would count as intrinsic reasons, if they existed. (This is not quite accurate when we are dealing with the phenomenon of a "moral authority" such as we suggested Kant might be on the question of suicide. The analogous point there is that we cannot identify someone as a moral authority, whose pronouncements we will take as indicators about what counts as an intrinsic reason, without having *some* prior views of our own concerning the identification of *some* facts as intrinsic reasons.) Incidentally, even though indicative reasons are logically secondary in the way I have just described, this does not undermine my claim to be distinguishing between indica-

tive and intrinsic force within the class of first-order reasons. Indicative reasons like the canvasser's credentials, even though they are logically secondary, are not at all like exclusionary reasons, or other reasons which are second-order in the sense that they govern whether or how we should act on first-order reasons which are assumed to exist.

The second previously-encountered point follows from the first. Because indicative reasons are logically secondary, they operate differently from intrinsic reasons in practical reasoning. What counts as an indicative reason depends in part on the state of the agent's relevant background knowledge. That in turn means that indicative reasons are "defeasible" in a way intrinsic reasons are not; they are subject to being completely deprived of force by the acquisition of new information.

Finally, the new point. It is important to distinguish between intrinsic and indicative reasons because phrases like "there is a moral obligation to . . ." or "there is moral reason to . . ." or "one ought to . . ." all create intensional contexts – that is, contexts in which substitution of materially equivalent descriptive phrases does not necessarily preserve truth value. It will be useful to consider an example other than the canvasser example. Imagine a situation where it is true of some possible act both (1) that by doing the act I can relieve substantial suffering at trivial cost to myself, and (2) that by doing the act, I will in some other respect make my friend Tom mildly unhappy. Clearly, I have a moral reason to do the act, even a moral obligation (in the common sense in which an obligation is just a significant "ought" but not necessarily any special kind of ought). Furthermore, because there is an act which I have a moral reason, or obligation, to do, and of which it is also true that it will make Tom unhappy, we might well say (in contexts where there was no danger of being misunderstood) that I have reason, or an obligation, to make Tom unhappy. But we certainly would not say this if we were trying to increase our philosophical understanding of the source and nature of moral reasons or obligations. In such a context, where we should be speaking precisely, to say I had reason, or an obligation, to make Tom unhappy, would be to say that the prospect of making Tom unhappy was itself something that counted in favor of doing the act. This is obviously false. So, even though "relieving substantial suffering" and "making Tom unhappy" are, in the present context, materially equivalent descriptions of a possible act of mine, it is true that I have reason to relieve suffering and false, strictly speaking, that I have reason to make Tom unhappy. I trust the reader will have noticed that the description that makes a strictly true proposition when it is inserted into the context "I have reason to . . ." is the description that refers to the feature of the act that constitutes an *intrinsic* moral reason for action.

Of course, in the example just discussed, there was no reason to think that the feature of the act "that it would make Tom unhappy" was even an indicative reason for action. But imagine now that it is. Suppose it just happens that my doing what I have moral reason to do almost invariably has the side effect of making Tom unhappy. If the connection is statistically strong enough, I might even be able to use Tom's prospective unhappiness as a guide to action. It would be an indicative reason. If Tom's prospective unhappiness is a useful indicator of what I have moral reason to do, then there is now a sense in which I have reason to do what

would make Tom unhappy, in a way that was not true before. As we have seen, even an indicative reason is a guiding reason. But if we are trying to understand the ultimate source and nature of our moral obligations, it would *still* be a mistake to claim that I have reason to do what would make Tom unhappy in the same sense in which I have reason to relieve suffering. Even if, in the present case, the act which would relieve suffering and the act which would make Tom unhappy are one and the same, the connection is only contingent. Tom's unhappiness might fail as an indicator in some other case. And even if it never failed, the fundamental reason why I should do the act would not be that it would make Tom unhappy, but that it would relieve suffering.

We see that even when we are considering materially equivalent descriptions, one of which refers to a feature which constitutes an indicative reason, and the other of which refers to a feature which constitutes an intrinsic reason, it is only the description in terms of the intrinsic reason of which we can truly say "there is moral reason to . . ." in the strictest sense. If we want to understand the real source and nature of our moral obligations, we must distinguish between indicative and intrinsic reasons.

Incidentally, although I cannot develop the argument here, I think it follows from what I have just been saying that the intrinsic/indicative distinction is the key to understanding how an agent should properly deal with those problems of practical application (and certain related supposed objections to consequentialism) that so-called "two-level" moral theories are designed to deal with. My discussion elsewhere of "indicator-rules" represents a beginning on just one aspect of a family of such issues.¹³

(Let me now add two quick final observations about the intrinsic/indicative distinction in general – observations that will tie up loose ends for some readers and seem quite unnecessary to others. (1) I have said earlier that giving to a canvasser with certain credentials is not what really matters morally, that it is only contingently connected with what does really matter, the relief of suffering. Someone might suggest that if we are evaluating the *character* of an agent, then whether she gave to a canvasser with particular credentials might well "really matter". Now certainly, different things matter in different contexts, and the context I have been focussing on is the context of decision-making, not of evaluation of character. But even in the latter context, it is not giving to people with certain credentials that *really* matters. For example, if for some reason the agent in question just did not understand what credentials of a certain sort normally signified, we might think her poorly socialized, but we would not regard her failure to give to bearers of those credentials as a moral failing. Even more strikingly, if the agent in question thought that, from her point of view as a decision-maker, giving to people with certain credentials *was* what really mattered in itself (as opposed to the relief of suffering), then we would not give her the same moral credit for giving to the canvasser as we would if she understood that the credentials were not what really mattered, but that they mattered only because of their connection with relieving suffering. In sum, what we want of the agent (in this context) is that she should respond properly to indicative reasons, which means being moved

13. Regan, *supra* note 1 at 1003-13.

by a correct understanding of their relation to the intrinsic reasons they indicate. (2) A totally different point – It might seem, given my emphasis on “what really matters” and the like, that the intrinsic/indicative distinction is of interest only to a moral realist like myself, and not to a conventionalist like, say, J.L. Mackie. But even a conventionalist ought to care about the distinction if he wants to speak consistently. Even if we merely *posit* what we shall treat as “mattering”, there is still a difference between what we have posited as mattering and what is merely a good indicator concerning what we have so posited. It is true there is a feature of the conventionalist approach which may seem to undercut the distinction. If we have posited that one thing matters, and if we notice or learn that some other thing is a sufficiently good indicator of the presence of that first thing, then, because we are not constrained in our positings of what matters by any beliefs about moral reality, we may be moved to posit that the second thing matters also, in the very same way in which the first thing matters. Even so, not every useful indicator of something that we have posited matters will be so tightly connected to that thing that we will be moved to posit that the indicator matters (in the same basic way) as well. And so, the distinction is still significant for any conventionalist whose conventionalism is not a mere excuse for sloppy argument.)

Authoritative Directives and Obedience

(With a Digression on “The Moral Obligation To Obey the Law”)

So much for general discussion of intrinsic and indicative reasons. Earlier I said that the central difference between Raz and me (concerning the operation of authority) is that I think authoritative directives are merely indicative reasons, and he thinks they are intrinsic reasons. In my view, our reason for following authoritative directives (when we have reason to) is that the authority, primarily because of superior factual knowledge, is better at identifying the acts we ought to do than we are ourselves. We use authoritative directives as indicators about the existence of underlying intrinsic reasons for action. Indeed, we often use them as indicators about the *balance* of intrinsic reasons. We do this, and more, if we adopt what I have elsewhere described as an “indicator-rule” of following authority, which also involves giving authoritative directives a degree of indicative second-order force (exclusionary force of a certain sort). But the force is always merely indicative.¹⁴ (Remember that in saying authoritative directives are merely indicative reasons, I am still talking only about cases which do not involve coordination problems. Where co-ordination is involved, directives may do more than indicate – they may create intrinsic reasons – but still without being intrinsic reasons themselves.)¹⁵ Now, Raz would not deny that authoritative directives are indicative reasons in the broad (non-exclusive) sense of the term. But he would deny that authoritative directives are *merely* indicative reasons (even in non-coordination contexts). For him, authoritative directives are intrinsic reasons also.

As I have mentioned previously, Raz never says in completely unambiguous terms that authoritative directives are intrinsic reasons. So I must now justify my claim that he is committed to this belief by his talk about obedience to authority.

14. *Ibid.*, 1003-18.

15. *Ibid.*, 1019-31.

What does it mean to obey an authority's directive? The first point, on which I am sure Raz and I agree, is that obedience involves something more than mere behavioral compliance with the directive. I cannot be said to obey in the fullest (and philosophically interesting) sense unless I do what I have been told to do *because I have been told to do it*. Consider an example. If a governmental authority commands me to avoid the use of cocaine, and if I subsequently avoid the use of cocaine, but for reasons which make no reference to the authority's command (I might even be a philosophical anarchist and regard the command, in itself, as a reason *in favor of* using cocaine), then I do not obey the authority in any interesting sense. My behavior is in compliance, but I do not "obey".

As I say, I am sure Raz agrees so far. Even so, let me address briefly two possible objections to the claim that obedience involves more than behavioral compliance. First, it might be objected that if I never use cocaine, I can hardly be said to *disobey* the authority's command. That is true. Behavioral non-compliance is a necessary condition for disobedience. But it does not follow that behavioral compliance is a sufficient condition for obedience. That would follow if obedience and disobedience were contradictories. But what the cocaine example shows is that they are not contradictories. They are merely contraries. There are cases, like the case where I abstain from cocaine for reasons totally independent of the authority's command, where I neither obey in the fullest sense nor disobey.

Second, it might be objected that we plainly do sometimes use the word "obey" to refer to nothing more than behavioral compliance. I admit that this is true. (That is why I have occasionally referred to obeying "in the fullest sense", or variants thereof.) I will even admit that people often appear to mean by "obedience" nothing more than behavioral compliance when they are discussing "the moral obligation to obey the law", which is the context where obedience is most likely to be the subject of philosophical discussion. Actually, there are specific and interesting reasons, which I shall explain shortly, why discussions of "the moral obligation to obey the law" tend to gloss over the difference between obedience in the strict sense and mere behavioral compliance. But even if the use of "obedience" to mean mere compliance turns out to be statistically the commonest use, the cocaine example ought to persuade the reader that ordinary usage also authorizes a strict sense of "obedience", which includes a motivational component. I am not just making it up.

I have said that Raz and I are both interested in the strict sense of "obey". We agree that for our purposes, I obey an authority only if I do what I have been told to do because I have been told to do it. But the phrase "because I have been told to do it" is still (multiply) ambiguous. Altering the original cocaine example a bit, let us now suppose that I regard government as better informed than I about the effects of cocaine. I therefore treat the directive not to use cocaine as an indicative reason for not using cocaine, and I do not use it.

In this situation, I not only avoid cocaine, but I avoid cocaine because I have been told to avoid it. Even so, if I regard the directive as a *merely* indicative reason, then I am still not obeying the directive when I rely on it. The reason is that I regard its force as provisional; if I somehow acquired evidence that the government was wrong about the effects of cocaine, I would regard the directive as deprived of all force as a guide to action. But I cannot be said to "obey" the directive if

I regard its significance as provisional in this way. If I am to obey, I must regard the directive, or complying with the directive, as something that matters *in itself*. In other words, I can only be said to obey the directive, in the strict sense, if I regard the fact that a certain act will comply with the directive as an intrinsic reason for action. (The importance of complying with the directive may of course depend on further facts about my relation to the authority, for example, facts which make complying with the directive an instance of discharging a duty of fair play or gratitude. But this is consistent with the claim that there is obedience only when compliance matters in itself. That some action will discharge a duty of fair play or gratitude, or whatever, may depend on many facts; but even so, that an action will discharge such a duty is a fact that matters in itself.)

If I regard the government's directive as a merely indicative reason, then even when I am guided by it, I am merely using it as evidence. This is true so long as I recognize the *possibility* that new evidence (about cocaine, not just about how the directive was adopted) would deprive the directive of its guiding force; it does not matter whether I actually discover any new evidence that has this effect. Using the directive as evidence is one way of "doing what one has been told to do because one has been told to do it". But it is not the way which constitutes obedience.¹⁶

Perhaps it is worth mentioning another case of "doing what one has been told to do because one has been told to do it" that also falls short of obedience – the case where one complies in order to set an example, hoping to encourage compliant behavior in others. If one is sufficiently Sidgwickian, one may even be trying to persuade others that authoritative directives are intrinsic reasons – one may be trying to induce *in others* an intention to "obey". But the person trying to set an example need not himself believe that the authoritative directives are intrinsic reasons for action; and if he does not, he does not obey, strictly speaking. Once again, he merely uses the authority, in a more or less devious way.

I have been arguing that an agent "obeys" authority, in the strict sense, only if she regards the authority's directives (or facts about what will constitute compliance with them) as intrinsic reasons for action. Raz might suggest that I have drawn the wrong conclusion from my examples. Raz might concede that one does not obey if one regards authoritative directives as merely indicative first-order reasons, but he might suggest that what obedience requires in addition is that the agent think of authoritative directives as *second-order* reasons, with exclusionary force. A full discussion of this suggestion would take us beyond the limits I have set for this essay, but I can indicate the main lines of my response quite briefly. I agree that authoritative directives are often second-order reasons with some sort of exclusionary force. But second-order force, and in particular exclusionary force, can be either intrinsic or indicative, just like first-order force.¹⁷ It seems to me, for reasons very similar to those I have given in connection with first-order reasons, that if we try to find the distinctive quality of obedience in the second-order force of the directives, then the second-order force must be intrinsic.

16. For a slightly more expansive presentation of this point, see the Appendix to "Authority and Value" *supra* note 1, especially pages 1094-5.

17. The discussion of "indicator-rules" in "Authority and Value", *ibid.*, is an elucidation of the kind of *indicative* exclusionary force I think authoritative directives often have.

sic. One way or another, thinking of ourselves as engaged in *obeying* authority requires us to think the authority's directives matter *in themselves*. If it is agreed that talk of obedience commits Raz to regarding authoritative directives as intrinsic reasons at some level (whether first-order or second-order), then what I shall say in a moment about the tensions this introduces into Raz's overall account applies in its essentials regardless of what level we focus on.

Before we go on, let me insert some remarks, previously promised, about the moral obligation to obey the law. These remarks are not immediately relevant to my disagreement with Raz, but I hope they may have enough interest in themselves to justify the divagation.

Consider again the case where there is a law forbidding the use of cocaine and where Jones abstains from cocaine, but for reasons that have nothing to do with the law. Plainly Jones's behavior is in compliance with the law. A further question is, Has Jones satisfied her moral obligation to obey this law (assuming there is such an obligation)? I think most people who believe there is a moral obligation to obey the law would think that Jones has satisfied her obligation to obey the law about cocaine. If I am right, then we see that what is referred to as "the moral obligation to obey the law" is normally thought to be satisfied by mere behavioral compliance with the law. (Notice one *could* believe that mere compliance does not satisfy the obligation – one could believe the obligation is satisfied only by obedience in the strong sense. But I think this is not the common view.)

This suggests two possibilities. Either, for most philosophers, "obedience" really means no more than behavioral compliance, or else the phrase "the moral obligation to obey the law" is a misnomer, and we should speak of "the moral obligation to comply with the law" instead. Now, I do not believe that most philosophers, if they thought about it, would equate obedience and mere behavioral compliance (although I believe that many philosophers have not considered what they mean by "obedience" as carefully as they should). I think the phrase "the moral obligation to obey the law" really is a misnomer, in the way I have suggested.

Is it possible that a phrase so entrenched in philosophers' talk as "the moral obligation to obey the law" should be a crude mistake? I think it is, but some explanation is obviously called for.

Some readers may have noticed, perhaps only half-consciously, that even though obedience (as I understand it) requires more than mere behavioral compliance, it somehow did not seem to make all that much difference when I suggested renaming "the moral obligation to obey the law" as "the moral obligation to comply with the law". Such readers have good antennae; they already have an inkling of the reason why a genuine misnomer has gone so long unnoticed.

The crucial point is this: Even if one's obligation in respect of the law is really only an obligation of behavioral compliance (that is, even if mere compliance is thought to satisfy the obligation), nonetheless, someone who believes in this obligation and is moved to action by it will not merely be complying with the law but will be *obeying* the law in the fullest sense. (I shall expand on why this is so in a moment.) Hence, even if some philosopher were careful to argue only for an obligation of behavioral compliance with the law, she would still in effect be trying to persuade people to obedience. If her arguments were effective, they

would produce belief in an obligation to comply, and they would bring about action on the basis of that belief, and that action would *be* not just compliance but obedience. (This would be so even if the philosopher cared neither about general knowledge of moral truth nor about obedience, but had as her sole purpose encouraging compliant behavior.)

Now, why do I say that someone who believes in a moral obligation to comply with the law and who is moved to action by that belief, in fact *obeys* the law? I say it because one who believes, in a philosophically sophisticated way, in a moral obligation to comply with the law believes, *ipso facto*, that the existence of the law is an intrinsic reason for doing what the law requires. (She may believe only that the law is a constituent element of an intrinsic reason for action, if she relies, for example, on something like the argument from gratitude or fair play. I shall let the statement that “the law is” an intrinsic reason include the possibility that the law is a constituent element of an intrinsic reason.)

There are two reasons why someone who believes in a moral obligation to comply with the law is committed to the belief that the law is an intrinsic reason for action – one reason is conceptual, and the other, we might say, is polemical. The conceptual reason we have already encountered in another context. We have seen that the phrase “there is a moral obligation to . . .” creates an intensional context, and that when we are speaking with philosophical rigor, we can only properly fill in the blank with a phrase that refers to an intrinsic reason for action. So, if there is something properly called “the moral obligation to comply with the law” in philosophical discussion, then the fact that an act will be in compliance with the law must be an intrinsic reason for doing the act. More briefly, the law must be an intrinsic reason.

As to the polemical reason why believers in a moral obligation to comply with the law must believe the law is an intrinsic reason for action: it is only if they believe this that there is any real controversy. Even among people (such as myself) who deny that there is a moral obligation to comply with the law, all would admit that many acts the law requires are acts we have a moral obligation to do, on some other ground. Similarly, most opponents of the moral obligation to comply with the law would admit that in many contexts the law is a useful indicator concerning acts we ought to do. In fact, once they have thought about the matter, opponents of the moral obligation to comply with the law ought also to admit that the law is frequently a *causally* relevant factor in determining what our specific moral obligations are. For example, the law often creates expectations in others about my behavior, which I am then morally required to live up to in order to satisfy my moral obligation to avoid doing harm. (Incidentally, it is not necessary, in order for the law to create such expectations, that it even be *thought* to be an intrinsic reason for action, though it is true that one way the law actually creates such expectations in some people is through their mistaken beliefs about the law’s moral significance.) Notice that even if the law contributes causally to determining our specific moral obligations, that does not make the law an intrinsic reason (or even a constituent element of an intrinsic reason). Causal connections are contingent; facts about what count as intrinsic reasons are not.¹⁸ Precisely because

18. Cf. Regan, *supra* note 1 at 1027-8.

none of the claims about the law that I have just listed is really controversial, none of them can be what proponents of a moral obligation to comply with the law are claiming. What proponents of such an obligation must be claiming – and what most of them do seem to be claiming even though they rarely distinguish explicitly between indicative and intrinsic reasons – is that the law is an intrinsic reason for action.

So, taking it as established that the proponent of the moral obligation to comply with the law regards the existence of the law as an intrinsic reason, let us return to the agent who truly believes (that is, who believes in a philosophically sophisticated way) in a moral obligation to comply with the law and who is moved to act by that belief. This agent does what the law requires, and she does what the law requires because she believes the existence of the law is an intrinsic reason for action. In other words, she *obeys* the law, in the fullest sense, just as I previously claimed. It follows, as I have noted previously, that a philosopher who argues for the existence of a moral obligation to comply with the law is arguing *in effect* for obedience, even though she may think the obligation is satisfied by mere compliance. Obedience is what her argument will produce if it succeeds, as an argument, in influencing action.

We now see why the phrase "the moral obligation to obey the law" has become entrenched, despite being a misnomer from the point of view of one who believes compliance satisfies the obligation. As a consequence, we see also why there is nothing in this usage of "obey" to refute my claim that true obedience is more than mere compliance – that it is compliance motivated by the belief that the law is an intrinsic reason for action.

Tensions in Raz's Position

So, Raz thinks authoritative directives are intrinsic reasons, and I think they are merely indicative reasons. We have at last got the central disagreement clearly before us. Who is right? The only way to settle this question is to see which view is part of the best analysis of how appeals to authority figure in sound practical reasoning. Raz has spelled out his analysis in considerable detail in many writings; and even I, who have given the issue much less attention, said more in my earlier article than I can repeat, or even effectively summarize, here. So there is no hope of settling the question here.

What I can do here is to point out some tensions in Raz's position which are introduced by his treating authoritative directives as intrinsic reasons. And I can then give some further reasons for thinking Raz has paid too little attention to the distinction between intrinsic and indicative reasons, and to what that distinction entails.

Let us start with the tensions in Raz's view. I shall limit myself to the two that seem most striking. First, there seems to be a lack of fit between Raz's normal justification thesis and the claim that authoritative directives are intrinsic reasons. The normal justification thesis says, roughly, that Jones (or any other agency capable of issuing directives) is an authority for Smith if Smith will do better at complying with those reasons for action which apply to him (independently of Jones's

directives) by doing what Jones says than by relying on his own judgment.¹⁹ The normal justification thesis is the linchpin of Raz's analysis of authority. (And lest the reader be in any doubt, I think the thesis is essentially correct.) But notice that it makes the question of who is an authority for Smith depend in part on *Smith's* capacities as a decision-maker – on Smith's informedness, his ability to conform his deliberation and behavior to what is required by genuine reasons, and so on. It seems to me there is something quite peculiar in first saying that Jones's authority over Smith depends on Smith's capacities (or really on his *in*-capacities, since it is only Smith's shortcomings that could make Jones a better guide), and then going on to say that *if* Jones in fact has authority, then his directives are intrinsic reasons for action, facts that matter morally *in themselves*. The way in which Smith's capacities figure in determining Jones's authority – and specifically the fact that if Smith's capacities are sufficiently strong then Jones become irrelevant – seems inconsistent with the idea that Jones's directives matter in themselves. It seems to make it all too clear that Jones's directives matter precisely as indicators of other facts. (Remember, even if Jones's directives are primarily indicators, they can still have second-order exclusionary force as well as first-order force. I myself think authoritative directives often have exclusionary force. But I think the exclusionary force, like the first-order force, is merely indicative and depends on the addressee's incapacities. Whatever sort of force we look at, first-order or second-order, it is hard to see how a directive which is without force when the addressee has sufficient capacity can be thought of as something that matters in itself.)²⁰

The second tension in Raz's view (actually a second manifestation of the same basic tension) appears at the point where Raz explicitly declines to take a position on whether one should obey an authoritative directive which is clearly wrong but wrong on a matter of little consequence.²¹ Normally, one cannot get oneself in trouble by taking no position, but here I think Raz does. If authoritative directives are really intrinsic reasons – facts which matter in themselves – then it ought to be obvious that one *should* obey even a clearly erroneous directive, provided the bad consequences of doing so are slight. After all, if the directive is an intrinsic reason, it must have some weight *of its own*. That weight will require obedience even in cases where the balance of all the other reasons clearly points the other way, provided the net balance pointing the other way is sufficiently small. It is easy to see why Raz recoils from accepting this argument and from saying that one should obey a directive which is clearly in error provided the error is of no great moment: saying that would fly in the face of the normal justification thesis, which depends on authority's being a reliable guide. (One should not follow even a generally reliable guide when it is *clearly* wrong.) But that just proves my point,

19. Raz, *supra* note 2 at 53-7.

20. I remind the reader yet again that I have excluded from the present discussion cases where authority promotes coordination. In such cases, an authoritative directive may be more than a mere indicator of other facts. But I would argue that it still gets its importance from its causal (not logical) connection with other facts, specifically facts about salience, which it may causally create. ("Authority and Value" *supra* note 1 at 1027-8) Even in the coordination context, there is an incongruity between the basic thrust of the normal justification thesis, according to which authority is a facilitator of behavior required by other reasons, and the claim that authoritative directives are intrinsic reasons, facts that matter in themselves.

21. Raz, *supra* note 2 at 62.

that the normal justification thesis does not fit with the view that authoritative directives are intrinsic reasons.²²

Raz tries to defend himself against these criticisms of mine by distinguishing between two projects he is engaged in – on the one hand, elucidating the common-sense concept of authority, and on the other hand, advancing substantive views about just who has what authority over whom.²³ But I do not think this distinction helps Raz. The aspects of Raz's position that create the present problems – the normal justification thesis, and the claim that authority is to be obeyed (which is what reveals that authoritative directives are intrinsic reasons), and the suspension of judgment about what to do in cases of clear but insignificant error – all appear in the chapters where Raz is elucidating the concept of authority.²⁴ So the tensions that I am pointing out are internal to Raz's analysis of the concept. Of course it might be that Raz is merely reproducing tensions which are part of the common-sense concept. But if that is so, it would argue against taking the common-sense concept as the basis for one's own further elaboration; and it certainly ought to be noted.

In any event, Raz's propounding of the normal justification thesis makes me doubt that he is actually elucidating the ordinary concept of authority. It is a consequence of the normal justification thesis, which Raz explicitly recognizes, that government is authoritative over different citizens with regard to different ranges of issues.²⁵ This is not the common-sense view as I understand it. I suppose Raz might say the common-sense view of authority includes the normal justification thesis, and that common sense simply has not noticed that this entails government has a patchwork authority. That seems to me implausible. The conclusion is just too obvious not to notice, given the premise. Instead, I think the common-sense view, of political authority at least, bases authority on tacit consent, or gratitude, or fair play, or democratic procedures, or what have you, and not on the normal justification thesis. Of course, any of these supposed common-sense grounds for authority might coexist with the truth of the normal justification thesis; but none of them is the same.

Actually, I should admit that sophisticated versions of some of these common-sense alternatives might also generate patchwork authority. For example, the argument from fair play might not cover someone who has been badly treated by society at every turn. But neither the justification for the patchiness, nor the pattern of it, will be the same as under the normal justification thesis. The normal justification thesis entails that one person may be less subject to political authority

22. Raz suggests that what he suspends judgment about in the passage I am discussing is whether a clear mistake puts a directive outside the "jurisdiction" of the authority. ("Facing Up", *supra* note 3 at 1184) Of course, a directive which is outside the authority's jurisdiction is not an intrinsic reason for action. And so, if it were uncertain whether there were jurisdiction, it would also be uncertain whether the directive was an intrinsic reason. But in fact there is no uncertainty here about jurisdiction in any ordinary sense. We do not normally think a mistake in judgment, even a clear mistake in judgment, undermines jurisdiction. And in the present context, what we are talking about is a situation where an authority makes a clear mistake in a particular case that is unambiguously the *sort* of case the authority normally decides better than the individual – the sort of case, in other words, where the normal justification thesis creates "jurisdiction", insofar as that idea is apposite. If the normal justification thesis ever justifies us in thinking that an authoritative directive is an intrinsic reason, it ought to do so here as well.

23. "Facing Up", *supra* note 3 at 1184-85.

24. Raz, *supra* note 2, chs. 2, 3.

25. Raz, *supra* note 2 at 74, 77-8, 100.

than another just because she is a better decision-maker. That is what I think common sense (mistakenly) rejects.

It is also worth pointing out that if I am right about the common-sense view, common sense can properly talk about obedience to authority in a way that Raz, if he is true to the normal justification thesis, cannot. Common sense can without inconsistency regard doing what is directed by an authority you have tacitly consented to, or doing what is required by gratitude, or whatever, as something that matters *in itself*. And it does so regard it.

Some Further Results of Slighting the Intrinsic/Indicative Distinction (With a Discussion of Arguments that Begin “It Would Be a Good Thing if...”)

Let me now turn to some other grounds for thinking Raz has paid too little attention to the distinction between intrinsic and indicative reasons.

At one point, Raz says that I misrepresent my own view when I claim it does not allow us to speak of “obeying” authority.²⁶ Raz says that I recommend the adoption of a rule of thumb whose content is “Obey the authority”, so that if we follow the rule of thumb, we do in fact obey the authority.²⁷ Raz makes two mistakes here. In the first place, I do not talk about rules of thumb, but about indicator-rules. I invented the phrase “indicator-rule”, and explained at some length what I meant by it,²⁸ precisely because I have come to believe that “rule of thumb” means different things to different people. And Raz at one point recognizes explicitly that his rules of thumb and my indicator-rules are not the same.²⁹

More importantly, the indicator-rule I recommend is not “Obey the authority” but “Follow the directives of the authority.”³⁰ The difference is significant. As I explained in the earlier article, if one adopts the indicator-rule “Follow the directives of the authority”, one will treat an authoritative directive as an indicative reason; one will treat it not merely as an indicator of the existence of certain intrinsic reasons, but as a reliable indicator, generally speaking, of the *balance* of intrinsic reasons; indeed, one will treat it as such a reliable indicator that one will normally be unwilling actively to seek or to evaluate other evidence concerning the balance of intrinsic reasons. (This last point is what makes me speak of an indicator-*rule*.) But for all of that, the directive is still merely an indicative reason (both with regard to its first-order and its second-order force). If, for example, evidence that the authority is mistaken on a particular occasion simply obtrudes itself upon one’s attention and is too persuasive to require further evaluation, then one will ignore the directive. The directive has no significance in itself. The directive is not an intrinsic reason (at any level), and adopting the indicator-rule “Follow the directives of the authority” does not mean we regard it as one.

It should be clear, in light of what I have previously said about obedience, why I do not state the indicator-rule as “Obey the authority.” I have argued that we

26. “Facing Up”, *supra* note 3 at 1185.

27. *Ibid.*, 1183, 1185.

28. Regan, *supra* note 1 at 1003-13.

29. “Facing Up”, *supra* note 3 at 1186, n. 70.

30. Regan, *supra* note 1 at 1013, 1015, 1019.

obey authority only if we regard the directives of authority as intrinsic reasons. There would therefore be a conceptual confusion in purporting to adopt an indicator-rule with the content "Obey the authority." That would require us to regard the directives of authority both as intrinsic reasons and as merely indicative reasons.

Another thing which suggests to me that Raz pays too little attention to the distinction between intrinsic and indicative reasons is his argument that consent can create an obligation to obey the law.³¹ (I shall say nothing in this essay about his more promising but still problematic argument concerning respect for law.)³²

Raz argues in effect as follows: Acts which manifest identification with one's reasonably just society are intrinsically valuable. If consent created a duty to obey authority, then people could manifest identification with their society by consenting to be bound by its authority. Therefore, if consent were binding (that is, if it created a duty to obey), people would have the opportunity to perform a particular sort of valuable act (giving binding consent). That would be a good thing. Therefore, consent is binding.

It is worth emphasizing that in this argument, the *conclusion* is that consent is binding. One way in which Raz improves on many other philosophers who would base obligation on consent is by seeing that the claim that consent is binding requires proof. Notice also that the crucial inference in Raz's argument is from "It would be a good thing if consent were binding" to "Consent is binding." The material about the possibility and the value of manifesting identification with one's society by consenting is there to support the intermediate conclusion that it would be a good thing if consent were binding.

In my article on Raz, I objected to this argument on the ground that the inference from "It would be a good thing if X" to "X" is fallacious.³³ Raz has responded by suggesting in effect that although this inference is not valid in general, it is valid when "X" has the form "There is a duty to do A."³⁴ In other words, Raz suggests we *can* infer from "It would be a good thing if there were a duty to do A" to "There is a duty to do A." I think Raz is still mistaken, and the most important part of the mistake can be explained by a failure to distinguish between intrinsic and indicative reasons.

Let us leave aside for a moment the general issue of inferences about duties, and let us look more closely at Raz's argument for the bindingness of consent. What Raz is trying to prove is that consent creates a duty to *obey* the law. If I am right about the nature of obedience, then what Raz is trying to prove is that consent to an authority constitutes an *intrinsic* reason for doing what the authority directs. His argument depends on the following inference: "It would be a good thing if consent constituted an intrinsic reason. Therefore consent does constitute an intrinsic reason." But it should be obvious that this inference is fallacious. Consent may or may not constitute an intrinsic reason; but whether it does so

31. Raz, *supra* note 2 at 88-94.

32. Raz's argument about respect for law is more promising because the version that appears in his response to me, ("Facing Up", 1196-97) much more than the version in his book, (*The Morality of Freedom*, 97, 99) suggests that it may be interpretable as an argument which is *not* of the form "It would be a good thing if...". The significance of this will become clear as I discuss Raz's argument about consent in the text. (The question about what Raz's argument actually is, is too complex to go into here.)

33. Regan, *supra* note 1 at 1036-39.

34. "Facing Up", *supra* note 3 at 1197.

is a *basic* moral fact. (I assume non-conventionalism; Raz is not a conventionalist.) To argue that consent constitutes an intrinsic reason on the ground that it would be a good thing if it did is as misguided as arguing for the truth of the Riemann Hypothesis in mathematics on the ground that it would be a good thing if it were true.

Some readers might think there is no analogy between Raz's argument for the bindingness of consent and my imagined argument for the Riemann Hypothesis on the ground that "It would be a good thing if . . .". Raz does at least attempt to move from a moral premise to a moral conclusion, not a mathematical one. But this is not enough to save Raz's procedure. The reason, to reiterate, is that the facts about what are intrinsic reasons are the basic moral facts. Some kinds of moral facts we can prove by arguments that begin "It would be a good thing if . . .". (I shall eventually consider an example of such a fact.) But not these basic facts. Indeed, any argument of the form "It would be a good thing if . . ." *presupposes* these basic moral facts. It is by reference to these facts that we identify what actions and consequences would count as "a good thing".³⁵

Now let us consider why Raz might have made his mistake. His argument fails as a proof that consent is an intrinsic reason. Would the same sort of argument fare better if we were trying to prove something was an *indicative* reason? Since consent to authority is not a very natural candidate for being an indicative reason, consider a different example, suggested by Raz's response to me.³⁶ Suppose we are thinking about giving to charity. Now, most plausible-looking charitable organizations have a reasonable record of using money given to them to relieve suffering; and it would certainly be a good thing if more people gave money to plausible-looking charitable organizations. That suggests as a possibility the following inference: "It would be a good thing if an organization's having plausible charitable credentials were an indicative reason for giving it money. Therefore, an organization's having plausible charitable credentials *is* an indicative reason for giving it money."

The inference just suggested is of the form "It would be a good thing if *X*; therefore, *X*" and it is much more plausible than the earlier inference about consent as an intrinsic reason. As it happens, I do not think even this latest inference is valid as stated. (I shall explain in a moment.) But a very similar inference, so similar that the reader may think the distinction hardly worth making, *is* valid. The valid inference is "It would be a good thing if people *acted on* an organization's having plausible charitable credentials as an indicative reason. Therefore, an organization's having plausible charitable credentials *is* an indicative reason." This inference is really nothing more than an application of the concept of an indicative reason: what it *is* for something to be an indicative reason is just for it to be true that that something reflects intrinsic reasons in such a way that it would be a good thing for people to act on the something as an indicative reason.

35. Notice that in objecting to the circularity of any attempt to prove consent is an intrinsic reason by an argument of the form "It would be a good thing if . . .", I am *not* objecting to the self-referential quality of Raz's notion of consent which he himself has pointed out (namely, that consent, if it works, is an act that creates an obligation by being intended to create an obligation). (*The Morality of Freedom*, 99) But even conceding that consent is not vitiated from the outset by being self-referential in this way, we still need an argument about why it actually *succeeds* in creating an obligation to obey, that is, about why, in this context at least, consent is an intrinsic reason.

36. "Facing Up", *supra* note 3 at 1197.

(I hope it is obvious that any apparent circularity here could be avoided. We could describe the way of acting on a consideration which I have referred to as "acting on it as an indicative reason" without using the phrase "indicative reason", and we could then define an indicative reason as a consideration that ought to be acted on that way. But if we wanted to be this picky, we should have started much earlier in the essay, when the expository complications were already great enough.)

Now I can explain why I rejected the previously suggested inference with the premise "It would be a good thing if an organization's having plausible charitable credentials were an indicative reason." What this premise really says, on my view, is that it would be a good thing if it were the case that organizations' having plausible charitable credentials were sufficiently reflective of their propensity to relieve suffering so that it would be a good thing if people gave to organizations with plausible charitable credentials. This premise is true. But of course this premise is perfectly consistent with the possibility that plausible credentials are a very *poor* indicator of a charitable organization's propensity for relieving suffering. The premise does not say plausible credentials are a good indicator; it just says it would be nice if they were (because good indicators are useful things to have). So this premise does not entail that plausible credentials actually are an indicative reason (which was the conclusion we proposed to derive from it).

Let us pause and take stock. We have found no valid inference of the form "It would be a good thing if X; therefore X." We have found an inference which can easily be mistaken for one of that form, and which can validly establish that something is an *indicative* reason. But we have not found any inference even remotely resembling one of this form which establishes that something is an *intrinsic* reason. And if we remember that the facts about what are intrinsic reasons are the basic moral facts, it should be clear that no such inference is going to be available to show that something is an intrinsic reason.

We may be able to reinforce some of our conclusions by considering an argument which is stated specifically in terms of the existence of a duty, and which Raz suggests that I, as a good consequentialist, should accept.³⁷ The argument goes: "It would be a good thing if people gave to charity. Therefore, there is a duty to give to charity." (Notice that "duty" here just refers to a significant moral ought, as it often does in common usage. There is no question of anything like an exclusionary reason for giving to charity.) This is a standard sort of consequentialist argument, and I do accept it – on a suitable interpretation. But there are ambiguities we must expose, both in the phrase "give to charity", and consequently in the sort of claim we are making when we say there is a "duty" to give to charity.

(I would also point out that this argument is *not* of the form "It would be a good thing if X; therefore X." Rather, it is of the form "It would be a good thing if Y; therefore X", where Y is "People give to charity", and X is "There is a duty to give to charity." Y and X are related in their content; but they are not the same. Nor, as we shall see presently, can we validly reformulate the argument so that it has the form "It would be a good thing if X; therefore X" by replacing the premise with "It would be a good thing if there were a duty to give to charity." It is actually

37. *Ibid.*

an argument with this latest premise that Raz states, with the shorter argument I principally focus on embedded in it.)

Now, as to the ambiguity of “give to charity”. When we talk about giving to charity, we are usually referring to giving to plausible-looking charitable organizations, with the assumption, but not the absolute assurance, that our contributions will be used to relieve suffering. Sometimes, however, we use the phrase “giving to charity” as if it were *synonymous* with relieving suffering (by a particular mechanism).

If we are treating “giving to charity” as synonymous with relieving suffering, then the argument we are considering can be restated as follows: “It would be a good thing if people relieved suffering; therefore, there is a duty to relieve suffering.” It seems to me that this is no argument at all (even though premise and conclusion are both true). To my mind, premise and conclusion are actually identical propositions. Each is merely a different way of saying “That an act is one of relieving suffering is an *intrinsic* reason for doing the act – the relief of suffering matters in itself.” That is what we believe about the relief of suffering, and therefore that is the most natural reading of both of these assertions about the relief of suffering.

I understand that some people may disagree with this collapsing of the argument. They may think there is some logical space between the premise and the conclusion. But people who take this view will see the space in the form of the question, “Why should any particular agent be moved by the impersonal goodness of the relief of suffering?” For reasons that go far beyond the scope of this essay, I think this is ultimately not an intelligible question. But in any event, if it is an intelligible question, it is clear that the argument we are considering simply begs it. It does nothing to answer it. Nor does Raz ever try to answer this question, which leads me to suspect that he also thinks it is ultimately not intelligible. (We obviously cannot deal with the question by saying “It would be a good thing if the agent were moved by this impersonal goodness,” since the precise issue is why even true claims of “it would be a good thing if . . .” matter.) In sum, if “giving to charity” is synonymous with relieving suffering (that is, if we assume in effect that when we give to a charitable organization we know with *certainly* that suffering will be relieved), then the argument we are considering is no argument at all. Either its premise and its conclusion are identical, or it is blatantly question-begging.

On the alternative interpretation of “giving to charity”, that phrase refers to giving to plausible-looking charitable organizations, with the assumption but without the absolute certainty that the money will relieve suffering. In our world, where most plausible-looking charitable organizations use a reasonable proportion of their receipts to relieve suffering, it is a good thing if people give to charity (in the present sense). People’s giving to charity will have the consequence – the *causal* consequence, not the logical consequence – that suffering is relieved. Therefore people do indeed have a duty to give to charity.

But notice that no one has a duty to give to a plausible-looking charitable organization if she just happens to know, for some special reason, that her contribution to this organization will for some reason *not* have the consequence of relieving suffering. In other words, *what we are saying* when we conclude, on the present interpretation of the argument, that people have a duty to give to charity, is just

that people ought to treat plausible-looking charitable credentials as an *indicative* reason for giving to an organization.

There is a genuine argument here. It does follow from the fact that in our world most plausible-looking charitable organizations are "for real" that one has a duty to give to such organizations. But the relevant sense of "duty" involves treating the plausible appearances as an *indicative* reason. Once again, we have found no argument with a premise of the form "it would be a good thing if . . ." that proves that anything is an *intrinsic* reason for action.

Two final observations, both concerning why we cannot replace the premise of the argument with "It would be a good thing if there were a duty to give to charity." First, depending on the sense of "giving to charity" we have in mind and the corresponding sense of "duty" which that makes relevant, this would make the premise effectively equivalent either to (1) "It would be a good thing if relieving suffering were an intrinsic reason" or to (2) "It would be a good thing if plausible charitable credentials were an indicative reason." We have already seen, in the passage *before* we started analyzing the argument "It would be a good thing if people gave to charity; therefore there is a duty to give to charity", that neither of these kinds of premise leads anywhere interesting. (Perhaps it is worth pointing out that the premise we had discovered earlier that *did* ground a valid argument – namely, "It would be a good thing if people *acted on* an organization's having plausible charitable credentials as an indicative reason" – is extensionally equivalent, though not quite intensionally equivalent, to "It would be a good thing if people gave to charity", on the interpretation of that premise which makes it ground a genuine argument – that is, with "giving to charity" interpreted to mean giving to organizations with plausible charitable credentials.)

The second reason why we cannot start with the premise "It would be a good thing if there were a duty to give to charity" may be more immediately intuitively revealing. Let us simply ask ourselves the question, "Why does the argument 'It would be a good thing if people gave to charity; therefore there is a duty to give to charity' work? What truth does this argument, on the proper interpretation, embody?" The truth it embodies is simply that people ought to do acts which can be expected to have good consequences. And the argument works because "giving to charity" is an act; it is something people can do. In contrast, "*there being a duty to give to charity*" is not an act; it is not something people can do. So the premise "It would be a good thing *if there were a duty to give to charity*" simply does not fit with the essential nature of the argument. It therefore leads nowhere. (If we were conventionalists, we might of course construct an argument about why people should *posit* a duty to give to charity. Positing a duty is something people could do. But even such an argument would not prove that there *was* such a duty. Whether there was such a duty would depend on whether people had done the relevant positing. And in any event, we are not conventionalists.) In short, once we see what truth it is that is embodied in the argument we are considering, we see why the argument cannot work with the premise "It would be a good thing if there were a duty . . .".

I hope it is now clear why we cannot prove that consent creates a duty to obey by any argument that begins "It would be a good thing if . . .". And I hope it is clear that what we *can* prove by such arguments are in effect propositions about

what constitute indicative reasons. We see again the importance of the intrinsic/indicative distinction.

Coda: The Case of the Disappearing Disagreement?

My reasons for distinguishing between intrinsic and indicative reasons go far beyond my disagreement with Raz or my concern with obedience. (For example, I have said that I think a proper appreciation of the intrinsic/indicative distinction is crucial to understanding how we really should accomplish what “two-level” moral theories are designed to accomplish.) But it may occur to the reader that if Raz rejected my analysis of obedience – and in particular if he rejected my claim that one obeys authority only when one regards the authority’s directives as intrinsic reasons for action – then much of our disagreement about authority and obedience would disappear.

That is true. And I would welcome progress toward agreement. But if Raz does not mean by “obedience” what I mean by it, then his usage is misleading, and he ought to give up the word. The philosophy of common sense is often careless about terms it employs freely; and it is careless about the meaning of “obedience”. But the meanings common sense vacillates between, I suggest, are (1) obedience as mere behavioral compliance and (2) obedience as treating authority’s directives as intrinsic reasons (or elements thereof). As I have said, I think Raz plainly has in mind more than (1). But if he thinks we obey authority if we regard authoritative directives as indicative reasons, or as triggers for indicator-rules (which is how I regard them), then he means less by “obedience” than common sense’s (2), and his usage is totally idiosyncratic.

My own belief is that Raz implicitly accepts my analysis of obedience and wants to insist that authoritative directives are intrinsic reasons for action. Unfortunately, I do not see any way to square this claim with the normal justification thesis – and in particular with the consequence of that thesis, that the normative force of an authoritative directive is conditional upon the addressee’s incapacity as a decision-maker.